DOCUMENT RESUME

ED 049 295 TM 000 482

AUTHOR Brown, Bobby R.; Lathrop, Robert L.
TITLE The Effects of Violations of Assumptions Upon Certain Tests of the Product Moment Correlation Coefficient.
PUB DATE Feb 71
NOTE 40p.; Paper presented at the Annual Meeting of the American Educational Research Association, New York, New York, February 1971

EDRS PRICE EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS *Correlation, *Research Methodology, *Sampling, *Statistical Analysis, *Tests of Significance

ABSTRACT

This study investigated the effects of non-normality of the marginal distributions of the bivariate surface upon the sampling distribution of certain tests of significance of differences for the product moment correlation coefficient. The effects of non-normality were found to be rather substantial and to be dependent upon: 1) the degree of correlaton in the population, 2) the types and extent of non-normality introduced, and 3) in some situations, the size of the samples drawn. A relationship between the variance of the sampling distributions of the test statistics and other effects of marginal non-normality was also observed. (Author)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

ED049295

TM 000 482

THE EFFECTS OF VIOLATIONS OF ASSUMPTIONS UPON CERTAIN TESTS OF THE PRODUCT MOMENT CORRELATION COEFFICIENT

Bobby R. Brown and Robert L. Lathrop

In the space of approximately 100 years, the study of joint variation or correlation of observables has developed from its rather crude beginnings as an awkward, tabular-aided, exploratory groping to the present level of sophistication evidenced in such statistical techniques as multiple, partial, and canonical correlation, factor analysis, and cluster analysis.

In spite of continued development and increasing popularity of more sophisticated statistical techniques, the product moment correlation coefficient (r) continues to be a useful and much used statistic. Two inferential tests concerning the population coefficient (ρ) based upon observed sample coefficients are often employed. These tests take the form H_0: ρ = (some hypothesized value between -1.0 and +1.0) for a single coefficient, and H_0: $\rho_1 = \rho_2$ for a pair of coefficients. The probabilistic accuracy of these tests requires the assumption of the normal bivariate model. While the assumptions required by the normal bivariate model may occasionally be met, there is reason to believe that much of the data of interest in the behavioral sciences may violate the required assumptions.

The purpose of this study was to investigate the effects of non-normality of the marginal distributions of the bivariate surface

upon the sampling distribution of certain tests of significance of differences for the product moment correlation coefficient.

While this study was primarily concerned with the effects of violation of assumptions upon test of r, a brief statement of the nature of the sampling distribution of r in the normal case may be helpful at this point.

The sampling distribution of r is dependent upon the value of ρ. For $\rho = 0$, the distribution of r is symmetrical. However, as ρ departs from 0, the sampling distribution of r becomes skewed, the skew becoming increasingly pronounced as ρ approaches unity.

The ρ-dependent sampling distribution of r complicates the task of devising tests of significance based on r. As the distribution of r becomes skewed, as estimate of the standard error based on a symmetrical distribution becomes increasingly less appropriate. A satisfactory solution to this problem demands either some method of adjusting the tests of r to account for the changing shape of the distribution of r or some method which normalizes the distribution of r across values of ρ.

"Student" (1908) first gave the sampling distribution of r for $\rho = 0$. Based on this work, it can be shown that when $\rho = 0$,

$$\frac{r}{\sqrt{(1 - r^2) / (N - 2)}}$$

is distributed as t with N - 2 degrees of freedom. This test, however, is of limited value as it only provides for a test of the significance of the departure of a single r from 0, and then only if $\rho = 0$ or small.

The distribution of r for $\rho \neq 0$ was obtained by Fisher (1915). He observed that "the curve of sampling of the correlation coefficient becomes extremely skewed toward the ends of its range, and in these regions changes in form rapidly as ρ is changed."

A paper, commonly referred to as the "Co-operative Study, by Soper, et al. (1916) greatly expanded upon Fisher's 1915 paper. This paper, an extensive theoretical investigation of the sampling distribution of r, provided the basis for several subsequent theoretical studies of r. One formula derived by Soper, et al., which gave the sampling distribution of r as a function of ρ, proved very useful in deriving the moments of r and aided in the development of David's (1938) tables of ordinates and areas of the distribution of r.

From theoretical studies of the distribution of r it is possible to specify the expected distribution of r for any value of ρ. However, this is not a satisfactory solution where the interest is in devising a test of r, because ρ is unknown.

Fisher's (1921) hyperbolic tangent transformation, hereafter the "r to z" transformation, provides the basis for tests of r by providing a transformed value for r which is nearly normal and almost independent of ρ in distribution. This transformation of r,

$$z = \tfrac{1}{2} \log_e \left[(1 + r) / (1 - r)\right],$$

has an expected value of z given approximately by

$$E(z) = \zeta = \tfrac{1}{2} \log_e \left[(1 + \rho) / (1 - \rho)\right].$$

The sampling variance of z is approximately

$$\sigma^2_z = \frac{1}{N - 3} .$$

Based on this transformation of r, it is possible to construct a test of significance of the departure of an obtained r from any hypothesized value of ρ (Fisher, 1950). This test statistic takes the form

$$\frac{z - \zeta}{\sqrt{1/(N - 3)}} ,$$

where ζ is the transformed value of the hypothesized ρ. The test statistic is referred to a table of the normal distribution to obtain the probability of a change departure as large as $z - \zeta$.

Using the "r to z" transformation it is also possible to test the hypothesis that two correlation coefficients were obtained from populations have the same ρ (Fisher, 1950). This test takes the form

$$\frac{z_1 - z_2}{\sqrt{[1/(N_1 - 3)] + [1/(N_2 - 3)]}}.$$

This test statistic is also referred to a table of the normal distribution.

The "r to z" transformation has certain other applications, such as allowing an average r to be computed from a series of r's; however, only the two tests described above are of interest in the present study.

Interest in the distribution of r for applied purposes began to decline following the introduction of the "r to z" transformation. However, two additional theoretical papers are of interest in the present study. Gayen (1951) derived the mathematical form of the distribution r and z from nonnormal populations. Gayen also pointed out an error in Fisher's (1921) derivation of the moments of the distribution of z. The correct expressions for the moments of z indicate that in the case of ρ not 0.0, the distribution of z is slightly less normal than had been supposed.

Of more importance, Gayen (1951) showed that while the shape of the distribution of z is not seriously affected by marginal nonnormality, "the variance of r is very sensitive to changes in the popu-

lation form." The effect of nonnormality upon the variance of *z* is of course as damaging to a test of *r* as would be departures of the *z* distribution from normality. It was also observed that the difference between the variance for normal populations and the variance of non-normal populations "diminishes gradually (though not very rapidly) as the sample size increases."

The results of Gayen's (1951) investigation suggest that a modification of the *z* transformation which would stabilize the variance would be a worthwhile improvement. Hotelling (1953) has proposed two modifications of the "*r* to *z*" transformation which he suggests should have more nearly constant variances than *z* and perhaps provide more nearly normal distributions as well. The expressions for the two modified transformations, termed z^* and z^{**}, are given below:

$$z^* = z - \frac{3z + r}{4n}$$

where $n = N - 1$, and

$$z^{**} = z - \frac{3z + r}{4n} - \frac{23z + 33r - 5r^3}{96n^2}$$

The expected values of z^* and z^{**} are:

$$E(z^*) = \zeta^* = \zeta - \frac{3\zeta + \rho}{4n}$$

$$E(z^{**}) = \zeta^{**} = \zeta - \frac{3\zeta + \rho}{4n} - \frac{23\zeta + 33\rho - 5\rho^3}{96n^2}.$$

The variance of z^* and z^{**} is $1/n$, where $1/n = 1/(N - 1)$.

The following empirical studies of the effects of violations of assumptions upon the distribution of *r* were reviewed: Baker, 1930; Pearson, 1931, 1932; Chesire *et al.*, 1932; Rider, 1932; Heath, 1961; Noris and Hjelm, 1961; Hjelm and Noris, 1962. Based on these prior

investigations, the following summary statements seem warranted:

1. For $\rho = .0$, the effects of marginal nonnormality are apparently minimal; as ρ departs from .0 the effects become more pronounced. Some early investigators seem to have over-estimated the insensitivity of the distribution of *r* to violations of assumptions due to their having employed $\rho = .0$ or small. Conclusions arrived at by Heath (1961) indicate that this over-generalization still occurs.

2. The "*r* to *z*" transformation is a very useful approximation, for normalization of the distribution of *r*. However, it is only an approximation, and as ρ departs from zero the approximation is less adequate.

3. Marginal nonnormality affects both the distribution and the variance of *z*, with the effects upon the variance being more pronounced. Disturbances in either the normality or the variance of *z* affect the tests of significance based on *z*.

4. Investigations of the effects of violations of assumptions on the distribution of *r* have not been as systematic as might be desired. This is unfortunate since it has been shown that several parameters interact to produce the observed effects. A systematic investigation of effects across (a) level of ρ, (b) several types and degrees of marginal nonnormality, and (c) a range of sample sizes would help provide a more unified picture of effects. Norris and Hjelm (1961), while providing the most systematic study encountered, investigated only two levels of ρ and gave no objective measures of

the degree of skewness or kurtosis (βs or γs)[1] in their populations.

5. Tests of goodness of fit have often been employed in empirical studies of the sampling distributions of r and z. These tests, while giving estimates of the overall fit of the obtained distributions with the expected distribution, may lead to underestimates of the observed departures from expected, especially in the tails of the distributions. The tails of the distributions are the chief areas of interest in relation ot significance tests.

6. The test of the significance of the difference between two r's,

$$\frac{z_1 - z_2}{\sigma(z_1 - z_2)},$$

seems not to have been subjected to empirical study. The effects of marginal nonnormality on this test are of practical interest.

7. Hotelling's (1953) adjustments of z (z* and z**) have not been empirically investigated. The effects of marginal nonnormality on these statistics are of practical interest, particularly if these adjustments should prove to be more accurate or less sensitive to nonnormality than z.

---

[1] In considering marginal nonnormality two parameters of skewness, β_1 and γ_1, and two parameters of kurtosis, β_2 and γ_2, are employed to specify the extent of nonnormality in the population distributions. The following relationships hold for β_1 and γ_1: $\gamma_1^2 = \beta_1$, and for β_2 and γ_2: $\gamma_2 = \beta_2 - 3.0$. For a normal distribution $\beta_1 = 0$ and $\beta_2 = 3.0$. Positive values of β_1 indicate either a positively or negatively skewed distribution. Values greater than 3.0 for β_2 indicate a leptokurtic distribution, while values less than 3.0 indicate a platykurtic distribution. Formulas for β_1 and β_2 are:

$$\beta_1 = \left[\frac{\Sigma x^3/N}{(\Sigma x^2/N)\ (\sqrt{\Sigma x^2/N})}\right]^2, \qquad \beta_2 = \frac{\Sigma x^4/N}{(\Sigma x^2/N)^2},$$

where $x = X - \bar{X}$.

The present investigations attempted to overcome some of the deficiencies and ommissions pointed out above.

PROCEDURE

All sampling experiments in this study were conducted on an IBM 360/67 computer at The Pennsylvania State University Computation Center. A Fortran IV computer program was written which permitted the generation of bivariate populations of specified size, correlation, and distribution characteristics. The effects of departures from normality of the marginal distributions were investigated by introducing nonnormality to the distributions of the populations generated.

Bivariate populations of approximately 10,000 cases each were generated for 32 populations consisting of eight forms of marginal distributions (normal, platykurtic, slight leptokurtic, marked leptokurtic, slightly skewed platykurtic, slight skew, moderate skew, and extreme skew) across four levels of ρ (.00, .30, .70, and .90). 2000 samples of size 100, 40, 20 and 10 were drawn from each population.

A single program was written which permitted the generation of populations, as well as computation of population parameters, drawing of samples from the population, computation of test statistics, evaluation of the normal theory probability of each statistic, tabulation of the obtained frequency in the critical regions for each distribution, and computation of indices of distribution characteristics (mean, σ^2, β_1, and β_2) for the sampling distributions. The parameters, ρ, N, $\beta_{1(x)}$, $\beta_{1(y)}$, $\beta_{2(x)}$, and $\beta_{2(y)}$ for each of the 32 populations are given in Table 1.

TABLE 1

THE PARAMETERS, ρ, N, $\beta_{1(x)}$, $\beta_{2(x)}$, $\beta_{1(y)}$, AND $\beta_{2(y)}$ FOR THE 32 POPULATIONS FROM WHICH SAMPLES WERE DRAWN

Marginal Distributions	ρ	N	$\beta_{1(x)}$	$\beta_{2(x)}$	$\beta_{1(y)}$	$\beta_{2(y)}$
Normal	.93	9974	0.00	2.94	0.00	2.94
	.69	9960	0.00	2.94	0.00	2.94
	.31	9952	0.00	2.94	0.00	2.95
	.02	9952	0.00	2.92	0.00	2.94
Platykurtic	.93	9974	0.00	1.91	0.00	1.91
	.68	9960	0.00	1.90	0.00	1.90
	.31	9952	0.00	1.90	0.00	1.92
	.04	9952	0.00	1.90	0.00	1.92
Slight Leptokurtic	.91	9974	0.00	4.62	0.00	4.59
	.67	9960	0.00	4.64	0.00	4.61
	.34	9952	0.00	4.58	0.00	4.62
	.00	9962	0.00	4.66	0.00	4.61
Marked Leptokurtic	.91	9974	0.00	6.18	0.00	6.21
	.69	9960	0.00	6.31	0.00	6.31
	.32	9958	0.00	6.31	0.00	6.36
	.07	9942	0.00	6.25	0.00	6.28
Slight Skew (Platy.)	.93	9974	0.24	2.69	0.23	2.68
	.69	9960	0.21	2.66	0.22	2.66
	.31	9962	0.22	2.67	0.23	2.69
	.03	9952	0.22	2.65	0.22	2.66
Slight Skew	.93	9974	0.39	3.28	0.38	3.26
	.70	9960	0.35	3.23	0.36	3.23
	.33	9952	0.36	3.24	0.37	3.27
	.06	9952	0.36	3.22	0.36	3.23
Moderate Skew	.92	9974	1.08	3.62	1.07	3.60
	.68	9960	1.01	3.55	1.03	3.56
	.31	9952	1.03	3.57	1.04	3.59
	.06	9952	1.03	3.53	1.02	3.55
Extreme Skew	.86	9974	2.01	4.77	2.07	4.87
	.69	9958	2.06	4.88	2.07	4.87
	.31	9958	2.02	4.83	2.07	4.90
	.05	9952	2.03	4.83	2.04	4.84

The marginal distributions of the eight types of populations investigated are shown in Figures 1 though 8. As the plotting of these eight frequency distributions was accomplished from punched card output, the number of cases was reduced from approximately 10,000 per population to approximately 1,000 per population to expedite handling of the data. The curves have been smoothed, otherwise these distributions are not different from the marginal distributions of the populations.

In choosing the extent and type of marginal nonnormality, the overriding consideration was that the distributions chosen cover the range and type which might be expected to occur in educational and psychological data. The distributions employed in this study were chosen after a survey of data available to the author and after examination of the distribution types chosen for inclusion in other methodological studies (Norton, 1952; Games and Lucas, 1966).

All sampling from the populations was random with replacement. Two thousand samples of size 10,20,40, and 100 were drawn from each population. For each sample drawn the following statistics were computed: $\underline{x}$, $\underline{z}$, $\underline{z}^*$, and $\underline{z}^{**}$.

The following test statistics were calculated for each sample:

$$\frac{z - \zeta}{\sigma_{(z - \zeta)}},$$

$$\frac{z^* - \zeta^*}{\sigma_{(z^* - \zeta^*)}},$$

and

$$\frac{z^{**} - \zeta^{**}}{\sigma_{(z^{**} - \zeta^{**})}},$$

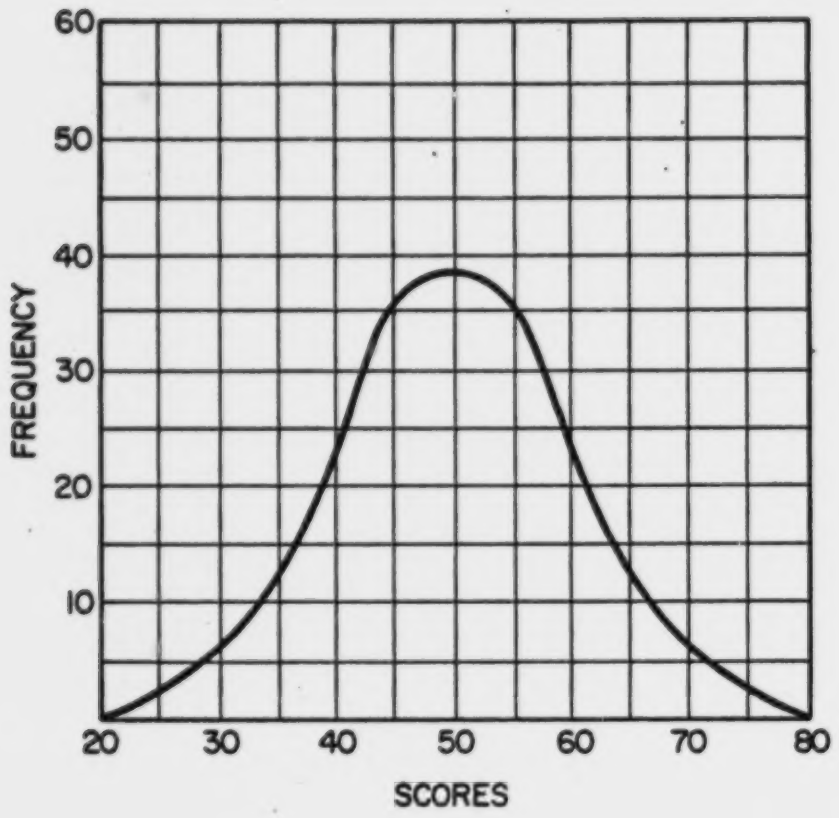

Figure 1. A Distribution Having the Characteristics of the X and Y Distributions in the Normal Populations ($\beta_1 = 0.00$, $\beta_2 = 2.94$).

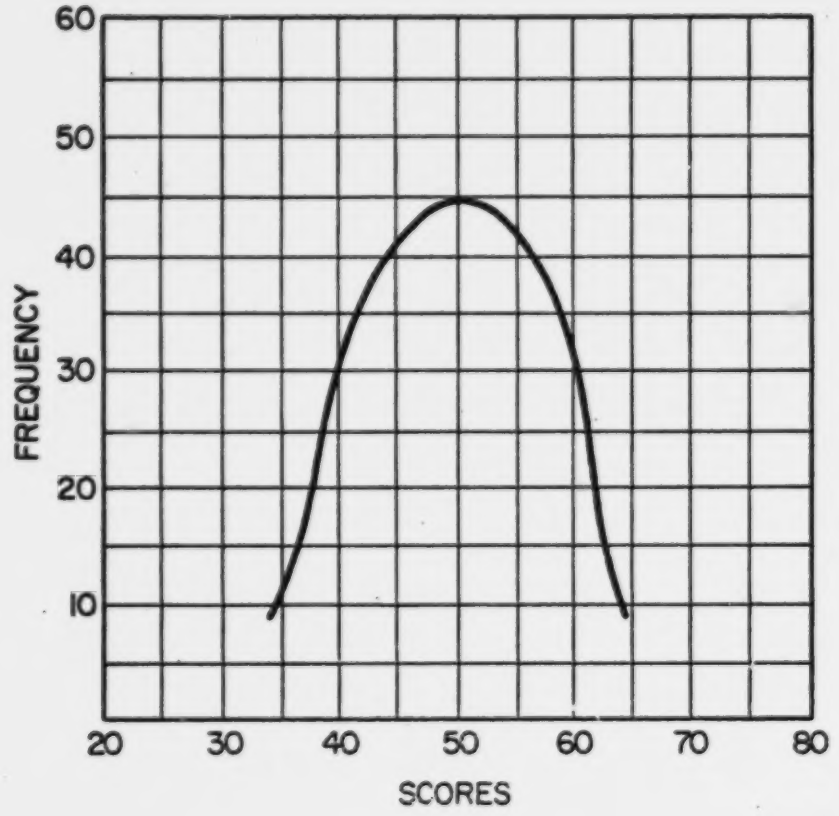

Figure 2. A Distribution Having the Characteristics of the X and Y Distributions in the Platykurtic Populations ($\beta_1 = 0.00$, $\beta_2 = 1.91$).

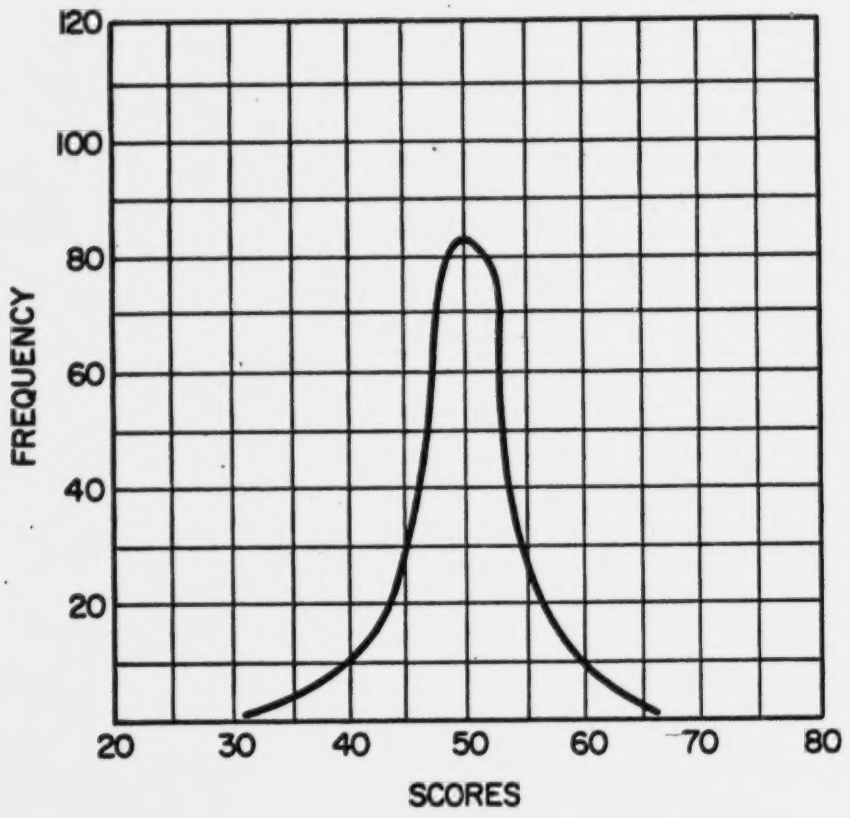

Figure 3. A Distribution Having the Characteristics of the X and Y Distributions in the Slight Leptokurtic Populations ($\beta_1 = 0.00$, $\beta_2 = 4.62$).

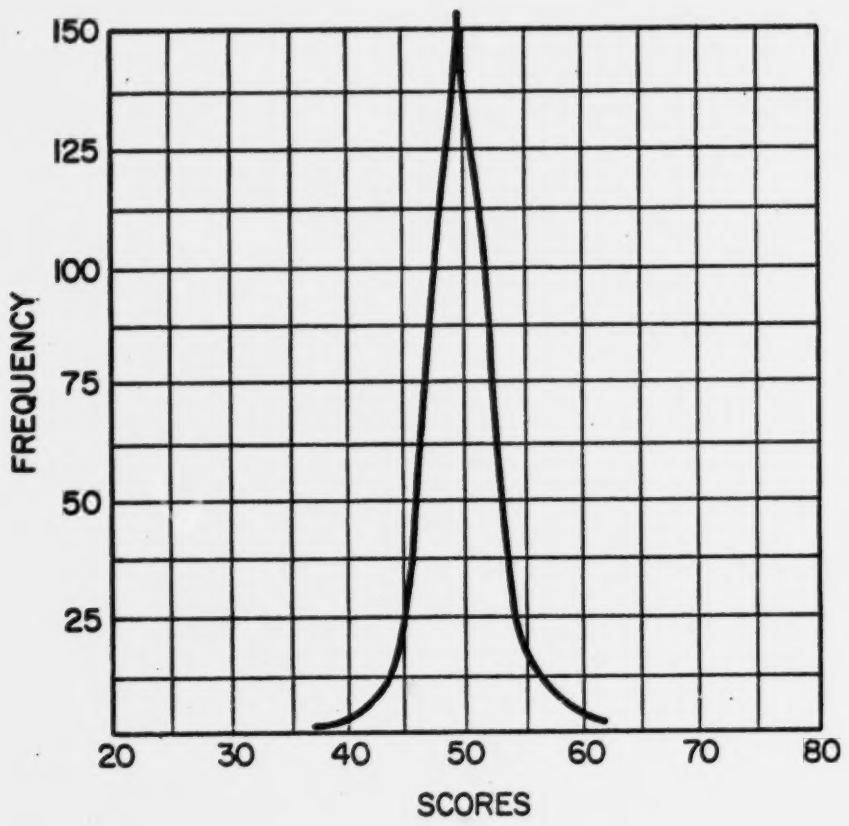

Figure 4. A Distribution Having the Characteristics of the X and Y Distributions in the Marked Leptokurtic Populations ($\beta_1 = 0.00$, $\beta_2 = 6.31$).

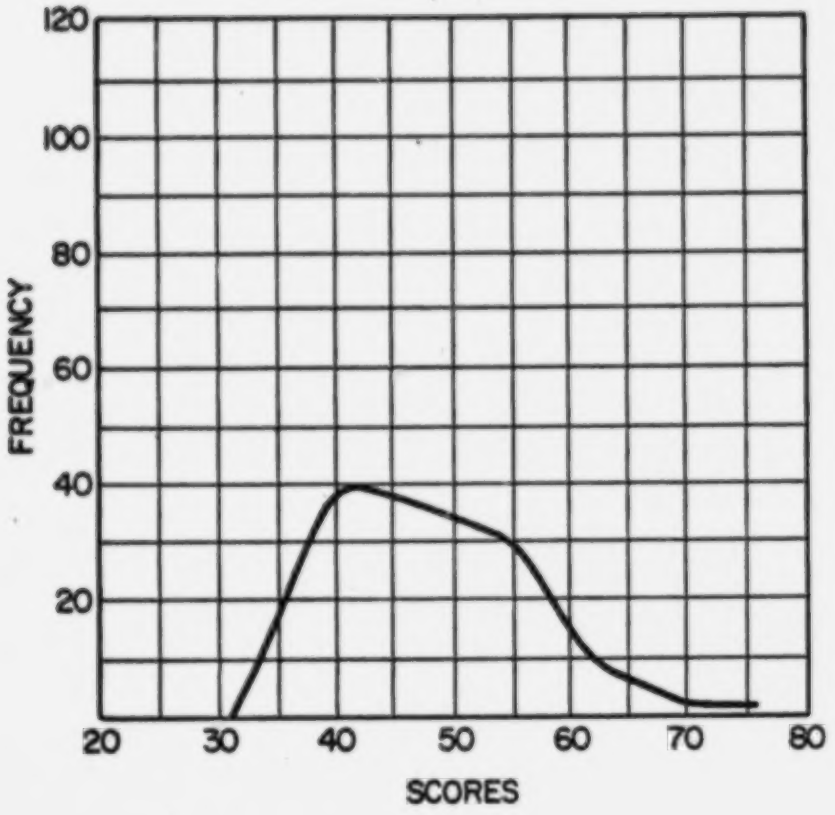

Figure 5. A Distribution Having the Characteristics of the X and Y Distributions in the Slight Skew (Platykurtic) Populations ($\beta_1 = 0.23$, $\beta_2 = 2.67$).

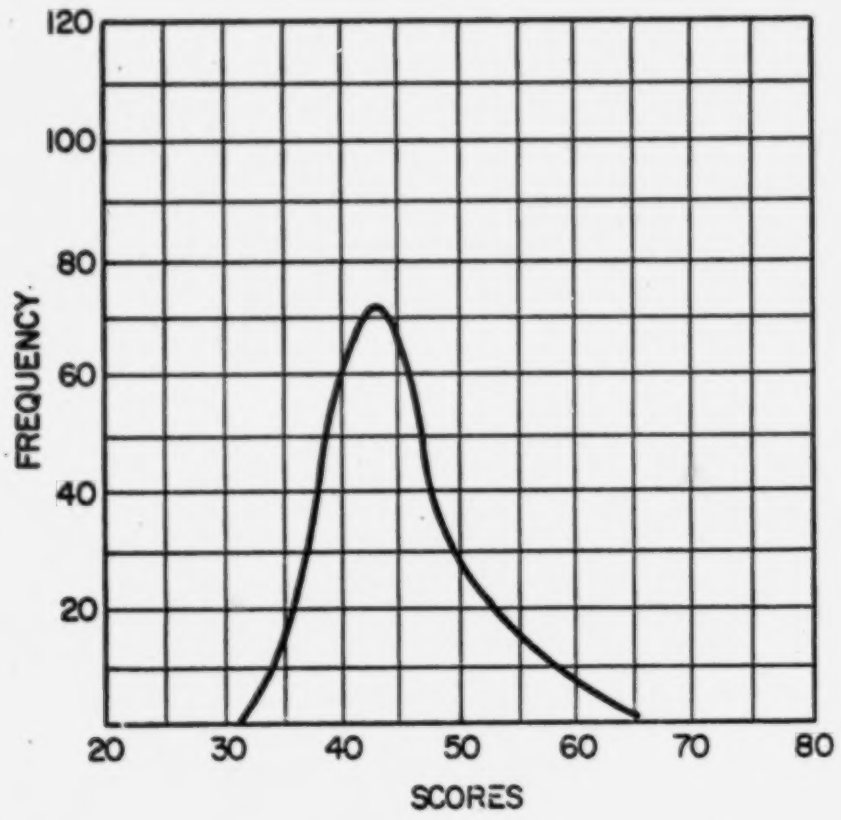

Figure 6. A Distribution Having the Characteristics of the X and Y Distributions in the Slight Skew Populations ($\beta_1 = 0.36$, $\beta_2 = 3.20$).

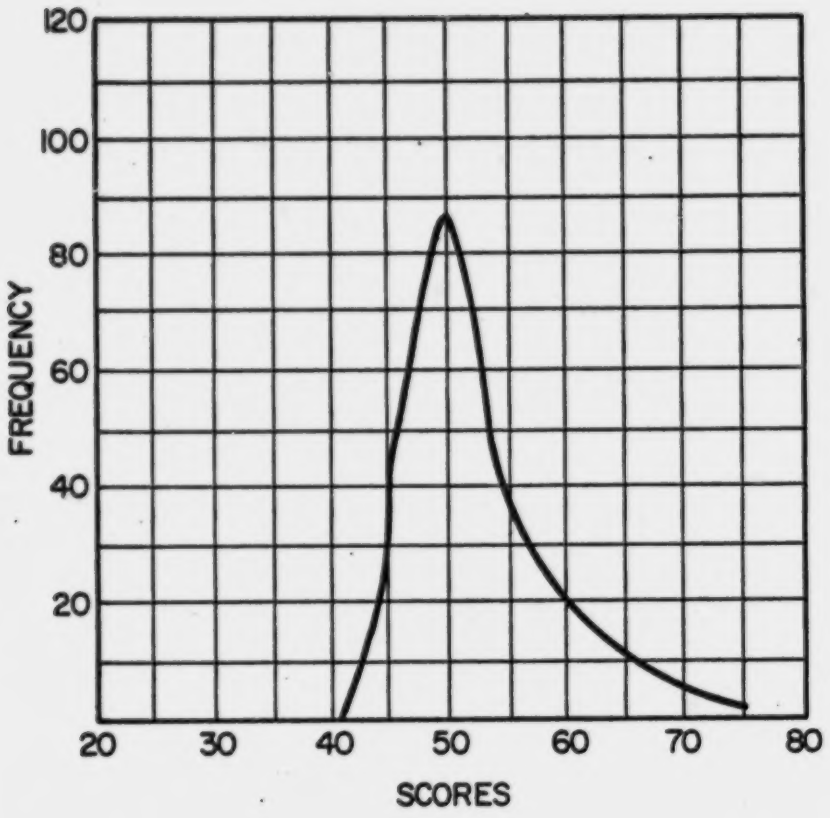

Figure 7. A Distribution Having the Characteristics of the X and Y Distributions in the Moderate Skew Populations ($\beta_1 = 1.05$, $\beta_2 = 3.58$).

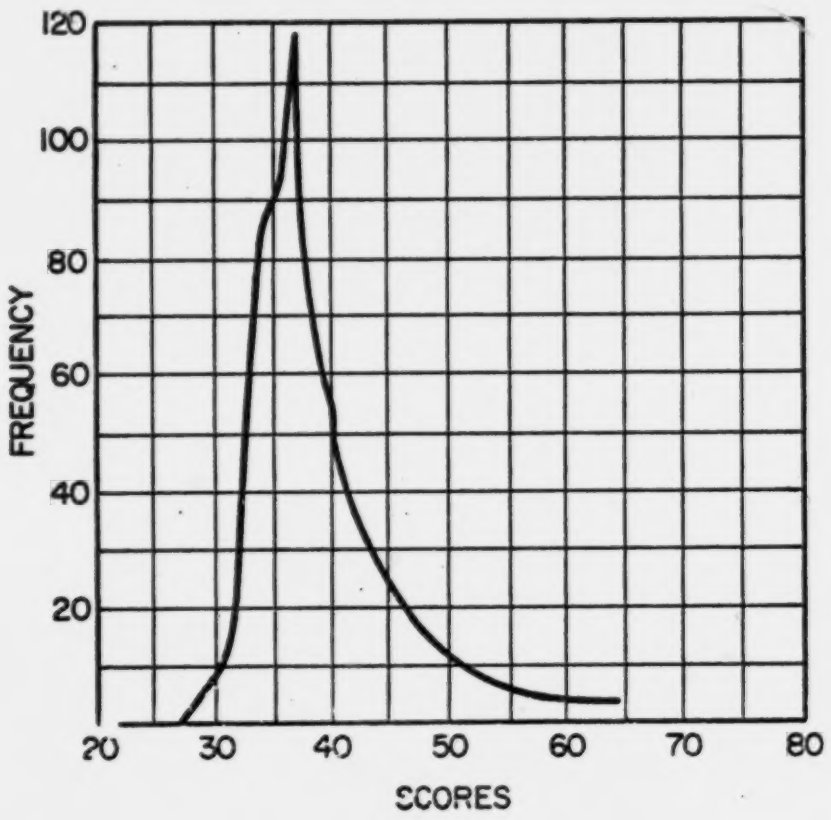

Figure 8. A Distribution Having the Characteristics of the X and Y Distributions in the Extreme Skew Populations ($\beta_1 = 2.06$, $\beta_2 = 4.67$).

where $z = \frac{1}{2} \log_e \left[(1 + r) / (1 - r) \right]$,

$$z^* = z - \frac{3z + r}{4n} - \frac{23z + 33r - 5r^3}{96n^2}$$

$$\zeta = \tfrac{1}{2}\log_e \left[(1 + \rho)/ (1 - \rho) \right],$$

$$\zeta^* = \zeta - \frac{3\zeta + \rho}{4n},$$

$$\zeta^{**} = \zeta - \frac{3\zeta + \rho}{4n} - \frac{23\zeta + 33\rho - 5\rho^3}{96n^2},$$

$$\sigma_{(z - \zeta)} = \sqrt{1/(N - 3)},$$

$$\sigma_{(z^* - \zeta^*)} = \sigma(z^{**} - \zeta^{**}) = \sqrt{1/(N - 1)},$$

and $n = N - 1$.

In each of these tests ρ was the actual population parameter previously computed. The normal theory probability of occurrence of each of the test statistics was evaluated by a normal probability density function, PRBZ, written for the System/360 by Knoble (1968). The probabilities returned by this function have an error less than .00000015 absolute. For each of these three tests performed, the frequency of observed probabilities was tabulated for the critical regions of .25,.125,.10, .05, .025, .01, and .005 for each tail, and of .50, .25, .20, .10, .05, .02, and .01 for both tails. The sample statistics from each of the above tests, as well as $\underline{z}$, $\underline{z}^*$, and $\underline{z}^{**}$ from each sample, were stored in arrays.

The following tests were performed on the statistics from successive pairs of samples:

$$\frac{z_1 - z_2}{\sigma_{(z_1 - z_2)}},$$

$$\frac{z^*_1 - z^*_2}{\sigma(z^*_1 - z^*_2)},$$

and $$\frac{z^{**}_1 - z^{**}_2}{\sigma(z^{**}_1 - z^{**}_2)},$$

where z, z^*, and z^{**} are as defined above, and where,

$$\sigma(z_1 - z_2) = \sqrt{[1/(N_1 - 3)] + [1/(N_2 - 3)]}$$

$$\sigma(z^*_1 - z^*_2) = \sigma(z^{**}_1 - z^{**}_2) = \sqrt{[1/(N_1 - 3)] + [1/(N_2 - 3)]}\ .$$

Evaluation of the normal probability of occurrence of each of these test statistics, tabulation of observed probabilities in the critical regions, and storage of the sample statistics were performed as described above. This procedure was repeated for each population and for each sample size (10,20,40, and 100) within each population.

At the completion of sampling from each population, sampling distributions of N = 2000 had been constructed for z, z^*, and z^{**}, as well as for the three tests of the form $z - \zeta$. As tests of the form $z_1 - z_2$ were performed for successive pairs, N = 1000 for their sampling distributions.

Taken all together there were nine sampling distributions: z, z^*, z^{**}, three of the form $z - \zeta$, and three of the form $z_1 - z_2$. For each of these distributions the following statistics were calculated: $\bar{x}$, σ^2, γ_1, γ_2, β_1, and β_2.

The computation of the distribution characteristic statistics for the sampling distributions completed one cycle of the program. The generation of all the data in the study can be thought of as nested cycles of the procedure described above. Within each population, sampling was cycled through four sizes of samples: 10,20,40, and 100. Within each population type, population generation was cycled through four

levels of ρ: approximately .00, .30, .70, and .90. And at the highest level, population type was cycled across eight types of populations: normal, platykurtic, slight leptokurtic, marked leptokurtic, slight skew (platykurtic), slight skew, moderate skew, and extreme skew.

To provide a check on the accuracy of all calculations and tabulations in the program, temporary print statements were inserted in the program and values were printed along with the actual cases sampled on one run of the program. The accuracy of all calculations was verified by hand calculation and by independent checks with other computer programs.

RESULTS

A complete detailed presentation of the results is beyond the scope of this paper. However, the presence of similarities among and patterns within the findings, make possible a comprehensive treatment of the findings of interest.

The results of the population sampling are presented in two major sections. In the first section results of tests of the form $\underline{z}_1 - \underline{z}_2$ are presented. The second section consists of the results of tests of the form $\underline{z} - \zeta$, along with data on the sampling distribution of $\underline{z}$, $\underline{z}^*$, and $\underline{z}^{**}$.

Tests of the Form $\underline{z}_1 - \underline{z}_2$

Tables 2 through 5 give the proportions observed in the critical regions for the .01, .05, .10, and .25 levels of significance for both tails of the sampling distributions for the test, $\underline{z}_1 - \underline{z}_2$. Only very small differences were found in the sampling distributions of tests of $\underline{z}_1 - \underline{z}_2$ and tests of $\underline{z}^* - \underline{z}^*_2$. With a few minor exceptions the obtained frequencies for $\underline{z}^*_1 - \underline{z}^*_2$ and $\underline{z}_1^{**} - \underline{z}^{**}_2$ are identical.

Figure 9 consists of a plot of the proportions observed in the tails of the sampling distributions of $\underline{z}_1 - \underline{z}_2$ at the .05 level of significance, as given in Table 2. The population types are laid out along the abscissa, and the results are plotted by sample size. Several results which appear throughout the data can be observed in this figure. Beginning at the far left of the abscissa, it can be seen that the obtained proportions for the normal populations, while not precisely equal to the expected proportions,

TABLE 2

PROPORTION OF CASES OBSERVED IN THE CRITICAL REGIONS, BASED ON 1000 TESTS OF z_1-z_2, FOR 2000 SAMPLES FROM EACH POPULATION TYPE, FOR SAMPLES OF SIZE 100, 40, 20, AND 10 (ρ APPROXIMATELY .90).

Marginal Distribution	N	ρ	Sig. Level: Two-tailed tests .01	.05	.10	.25
Normal	100	.93	.008	.038	.088	.208
Platykurtic		.93	.018	.078	.135	.278
Slight Leptokurtic		.91	.028	.111	.186	.350
Marked Leptokurtic		.91	.098	.197	.291	.451
Slight Skew (Platy.)		.93	.017	.054	.104	.249
Slight Skew		.93	.016	.068	.130	.284
Moderate Skew		.92	.023	.097	.163	.336
Extreme Skew		.86	.042	.127	.199	.352
Normal	40	.93	.013	.051	.092	.235
Platykurtic		.93	.017	.066	.112	.276
Slight Leptokurtic		.91	.056	.137	.207	.358
Marked Leptokurtic		.91	.086	.181	.280	.428
Slight Skew (Platy.)		.93	.014	.062	.112	.266
Slight Skew		.93	.019	.064	.124	.281
Moderate Skew		.92	.039	.110	.179	.343
Extreme Skew		.86	.061	.158	.237	.407
Normal	20	.93	.013	.052	.094	.242
Platykurtic		.93	.012	.061	.120	.278
Slight Leptokurtic		.91	.042	.119	.184	.331
Marked Leptokurtic		.91	.080	.167	.252	.422
Slight Skew (Platy.)		.93	.015	.049	.098	.254
Slight Skew		.93	.015	.055	.106	.263
Moderate Skew		.92	.031	.099	.163	.335
Extreme Skew		.86	.055	.128	.200	.366
Normal	10	.93	.016	.052	.097	.248
Platykurtic		.93	.021	.065	.117	.249
Slight Leptokurtic		.91	.083	.153	.207	.379
Marked Leptokurtic		.91	.086	.178	.261	.398
Slight Skew (Platy.)		.93	.022	.053	.091	.218
Slight Skew		.93	.023	.054	.096	.227
Moderate Skew		.92	.045	.104	.167	.327
Extreme Skew		.86	.071	.154	.236	.392

TABLE 3

PROPORTION OF CASES OBSERVED IN THE CRITICAL REGIONS, BASED ON 1000 TESTS OF z_1-z_2, FOR 2000 SAMPLES FROM EACH POPULATION TYPE, FOR SAMPLES OF SIZE 100, 40, 20, AND 10 (ρ APPROXIMATELY .70).

Marginal Distribution	N	ρ	Sig. Level: Two-tailed tests .01	.05	.10	.25
Normal	100	.69	.004	.040	.092	.229
Platykurtic		.68	.012	.051	.108	.241
Slight Leptokurtic		.67	.013	.050	.122	.271
Marked Leptokurtic		.69	.028	.086	.158	.323
Slight Skew (Platy.)		.69	.009	.046	.090	.252
Slight Skew		.70	.015	.047	.095	.250
Moderate Skew		.70	.023	.070	.135	.317
Extreme Skew		.69	.045	.133	.207	.399
Normal	40	.69	.008	.044	.100	.245
Platykurtic		.68	.013	.063	.107	.258
Slight Leptokurtic		.67	.015	.066	.125	.299
Marked Leptokurtic		.69	.033	.103	.177	.341
Slight Skew (Platy.)		.69	.012	.053	.103	.266
Slight Skew		.70	.014	.058	.099	.271
Moderate Skew		.70	.021	.082	.132	.320
Extreme Skew		.69	.048	.139	.227	.393
Normal	20	.69	.007	.050	.089	.244
Platykurtic		.68	.013	.061	.116	.257
Slight Leptokurtic		.67	.009	.061	.117	.270
Marked Leptokurtic		.69	.018	.073	.140	.322
Slight Skew (Platy.)		.69	.012	.052	.102	.243
Slight Skew		.70	.011	.058	.103	.246
Moderate Skew		.70	.024	.081	.145	.300
Extreme Skew		.69	.040	.125	.212	.373
Normal	10	.69	.012	.042	.087	.209
Platykurtic		.68	.017	.061	.117	.273
Slight Leptokurtic		.67	.016	.054	.112	.275
Marked Leptokurtic		.69	.023	.077	.140	.293
Slight Skew (Platy.)		.69	.009	.050	.113	.267
Slight Skew		.70	.010	.051	.108	.255
Moderate Skew		.70	.019	.082	.138	.314
Extreme Skew		.69	.043	.121	.191	.365

TABLE 4

PROPORTION OF CASES OBSERVED IN THE CRITICAL REGIONS, BASED ON 1000 TESTS OF z_1-z_2, FOR 2000 SAMPLES FROM EACH POPULATION TYPE, FOR SAMPLES OF SIZE 100, 40, 20, AND 10 (ρ APPROXIMATELY .30).

Marginal Distribution	N	ρ	Sig. Level: Two-tailed tests			
			.01	.05	.10	.25
Normal	100	.31	.010	.044	.088	.240
Platykurtic		.31	.011	.052	.098	.240
Slight Leptokurtic		.34	.007	.043	.090	.232
Marked Leptokurtic		.32	.014	.066	.106	.258
Slight Skew (Platy.)		.31	.012	.057	.102	.246
Slight Skew		.33	.015	.058	.110	.246
Moderate Skew		.31	.018	.069	.134	.268
Extreme Skew		.31	.041	.133	.204	.383
Normal	40	.31	.010	.039	.088	.236
Platykurtic		.31	.016	.054	.106	.244
Slight Leptokurtic		.34	.014	.055	.093	.254
Marked Leptokurtic		.32	.017	.056	.101	.251
Slight Skew (Platy.)		.31	.010	.054	.110	.245
Slight Skew		.33	.012	.053	.109	.246
Moderate Skew		.31	.011	.076	.129	.270
Extreme Skew		.31	.040	.119	.192	.351
Normal	20	.31	.004	.044	.098	.242
Platykurtic		.31	.006	.054	.106	.258
Slight Leptokurtic		.34	.010	.054	.111	.255
Marked Leptokurtic		.32	.013	.068	.125	.288
Slight Skew (Platy.)		.31	.008	.056	.108	.256
Slight Skew		.33	.007	.053	.112	.258
Moderate Skew		.31	.013	.058	.123	.278
Extreme Skew		.31	.020	.105	.179	.326
Normal	10	.31	.015	.055	.093	.217
Platykurtic		.31	.008	.048	.091	.218
Slight Leptokurtic		.34	.011	.044	.086	.225
Marked Leptokurtic		.32	.014	.058	.115	.247
Slight Skew (Platy.)		.31	.009	.041	.076	.230
Slight Skew		.33	.009	.039	.085	.224
Moderate Skew		.31	.011	.059	.094	.251
Extreme Skew		.31	.037	.124	.191	.338

TABLE 5

PROPORTION OF CASES OBSERVED IN THE CRITICAL REGIONS, BASED ON 1000 TESTS OF $\bar{x}_1 - \bar{x}_2$, FOR 2000 SAMPLES FROM EACH POPULATION TYPE, FOR SAMPLES OF SIZE 100, 40, 20, AND 10 (ρ APPROXIMATELY .00).

Marginal Distribution	N	ρ	Sig. Level: Two-tailed tests .01	.05	.10	.25
Normal	100	.02	.013	.041	.091	.245
Platykurtic		.04	.006	.048	.094	.246
Slight Leptokurtic		.00	.011	.040	.084	.247
Marked Leptokurtic		.07	.017	.061	.110	.243
Slight Skew (Platy.)		.03	.010	.047	.098	.232
Slight Skew		.06	.008	.049	.101	.245
Moderate Skew		.06	.010	.045	.105	.253
Extreme Skew		.05	.011	.052	.097	.228
Normal	40	.02	.009	.054	.109	.249
Platykurtic		.04	.011	.053	.121	.267
Slight Leptokurtic		.00	.011	.045	.108	.258
Marked Leptokurtic		.07	.022	.061	.106	.251
Slight Skew (Platy.)		.03	.015	.058	.115	.262
Slight Skew		.06	.018	.062	.113	.267
Moderate Skew		.06	.014	.067	.127	.259
Extreme Skew		.05	.009	.055	.101	.240
Normal	20	.02	.014	.055	.094	.238
Platykurtic		.04	.007	.051	.095	.240
Slight Leptokurtic		.00	.011	.043	.092	.251
Marked Leptokurtic		.07	.014	.060	.123	.278
Slight Skew (Platy.)		.03	.007	.052	.104	.233
Slight Skew		.06	.009	.047	.108	.240
Moderate Skew		.06	.007	.041	.092	.248
Extreme Skew		.05	.008	.063	.118	.240
Normal	10	.02	.010	.061	.108	.235
Platykurtic		.04	.009	.050	.084	.244
Slight Leptokurtic		.00	.012	.046	.100	.244
Marked Leptokurtic		.07	.014	.050	.105	.263
Slight Skew (Platy.)		.03	.006	.045	.084	.235
Slight Skew		.06	.009	.040	.086	.239
Moderate Skew		.06	.011	.045	.092	.250
Extreme Skew		.05	.011	.062	.104	.228

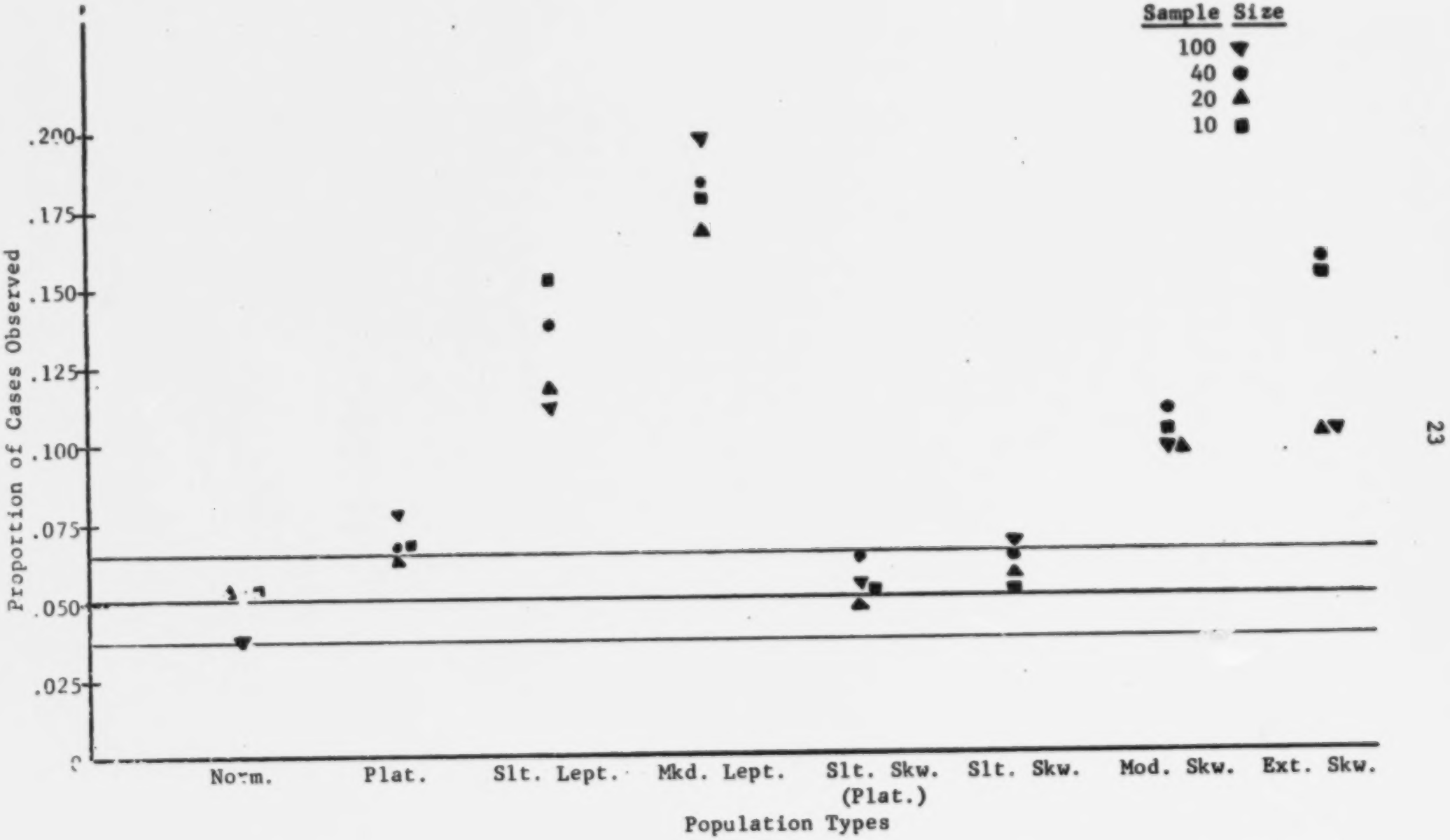

Figure 2. Proportions Observed in the Tails of the Sampling Distributions of $z_1 - z_2$ at the .05 Level of Significance for All Population Types, Across All Sample Sizes, for ρ Approximately .90.

23

are within the P = .95 confidence interval. With the exception of two populations, slight skew (platykurtic) and slight skew, all the nonnormal populations give rise to obtained proportions which exceed the expected. The marked leptokurtic and extreme skew populations give rise to rather extreme excess in the critical region.

From Figure 9 it can be seen that sample size does not appear to have any systematic effect on the excess observed. This same observation can be made from Figure 10. In Figure 10 proportions observed in the tails of the sampling distributions of $\underline{z}_1 - \underline{z}_2$ and $\underline{z}^*_1 - \underline{z}^*_2$ have been plotted from Tables 2 and 6. Sample size is given along the abscissa. The three populations plotted are normal, marked leptokurtic, and extreme skew. The rather pronounced excess in observed proportion is again seen for the nonnormal populations.

As can be seen from a comparison of proportions observed shown in Figure 10 there appears to be very little difference in the sampling distribution of tests of $\underline{z}1 - \underline{z}_2$ and tests of $\underline{z}^*_1 - \underline{z}^*_2$.

The relationship between ρ and the effects of marginal nonnormality can be seen in Figure 11. This figure shows the proportions observed across population types for ρ approximately equal to .00, .30, .70, and .90 for samples of size 100, at the .05 level of significance.

For $\rho = .00$ it can be seen that departures from the expected proportions are slight for all population types, and none departs from the confidence interval.

An interesting comparison between the effect of leptokurtic and skewed marginal distributions is possible from Figure 11. For moderate and extreme skew distributions departures are seen for ρ approximately equal to .90, indicate that ρ need not be large before the effects of skew are observed. For moderate and marked leptokurtic distributions, on the other hand, the effect, while great at ρ approximately equal to .90 is much less for ρ of .70

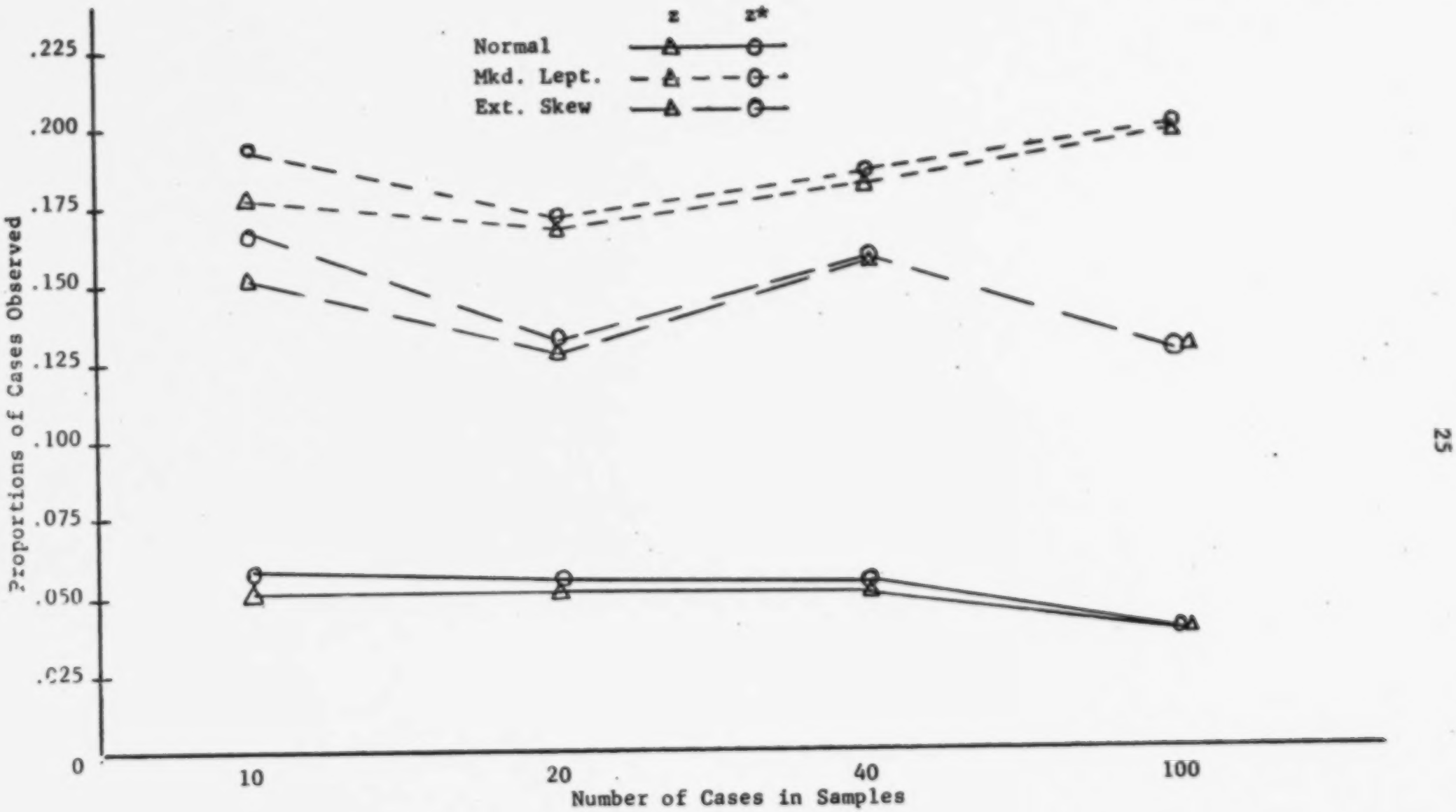

Figure 12. Proportions Observed in the Tails of the Sampling Distributions of $\underline{z}_1 - \underline{z}_2$ and $\underline{z}^*_1 - \underline{z}^*_2$ at the .05 Level of Significance for Normal, Marked Leptokurtic, and Extreme Skew Populations, Across All Sample Sizes, for ρ Approximately .90.

25

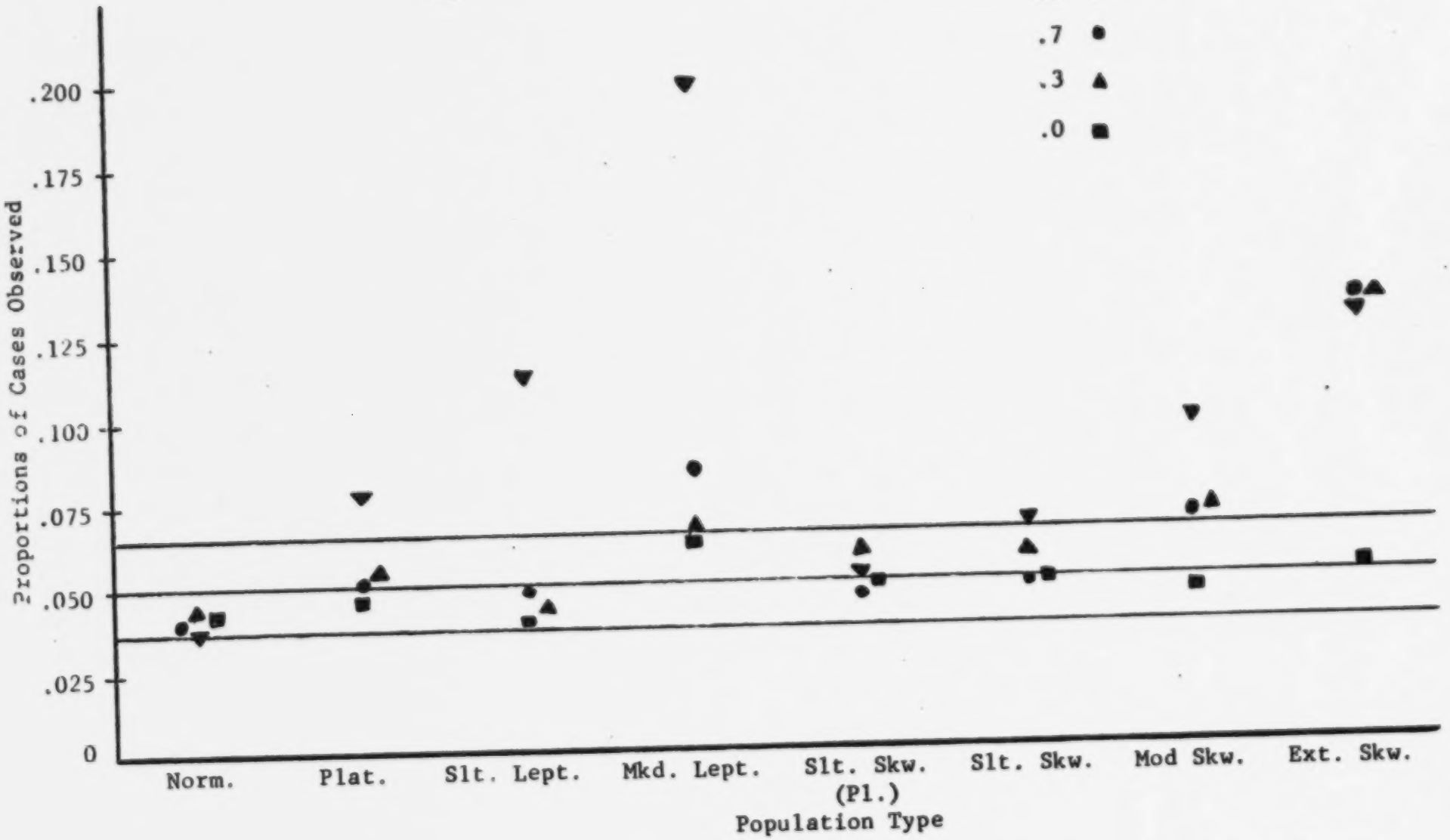

Figure 11. Proportions Observed in the Tails of the Sampling Distributions of $z_1 - z_2$ at the .05 Level of Significance for All Population Types, Across All Levels of ρ, for Sample Size 100.

and small for ρ of .30.

The expected variance of the sampling distributions for all the tests investigated in this study is 1.0. Table 6 shows the observed variance of the sampling distributions for $z_1 - z_2$, and $z^*_1 - z^*_2$ for all population types, across all levels of ρ, for samples of size 100. Figure 12 shows the observed variance of the test $z_1 - z_2$ across population types, plotted for ρ approximately equal to .00, .30, .70, and .90. In addition, the obtained proportions previously plotted in Figure 11 are also plotted in Figure 12. The scale for the obtained proportions is given on the left of the figure; the scale for the variance is given on the right of the figure.

As can be seen in Figure 12, there is a striking similarity between the relative excess in proportions observed and observed variance at any point on the figure. This similarity strongly suggests that the effects of nonnormality on tests of the form $z_1 - z_2$, seen in proportions obtained, is a result of departures of sampling distribution variance from the expected value.

Tests of the Form $z - \zeta$

We turn now to tests of obtained zs against ζ. In all cases ζ was based on ρ of the population being sampled. The results previously noted for σ^2, ρ, N, and population type were not found to be appreciably different for tests of the form $z - \zeta$. However, due to the fact that the tails of the distributions can be considered separately for these test one additional findings of interest was obtained. In Figures 13 and 14 proportions observed have been given for the upper (ζ less than z) and lower (ζ greater than z) tails instead of the sum of the proportions for both tails. It can readily be seen from both figures that the proportions observed are not equally distributed in the two tails.

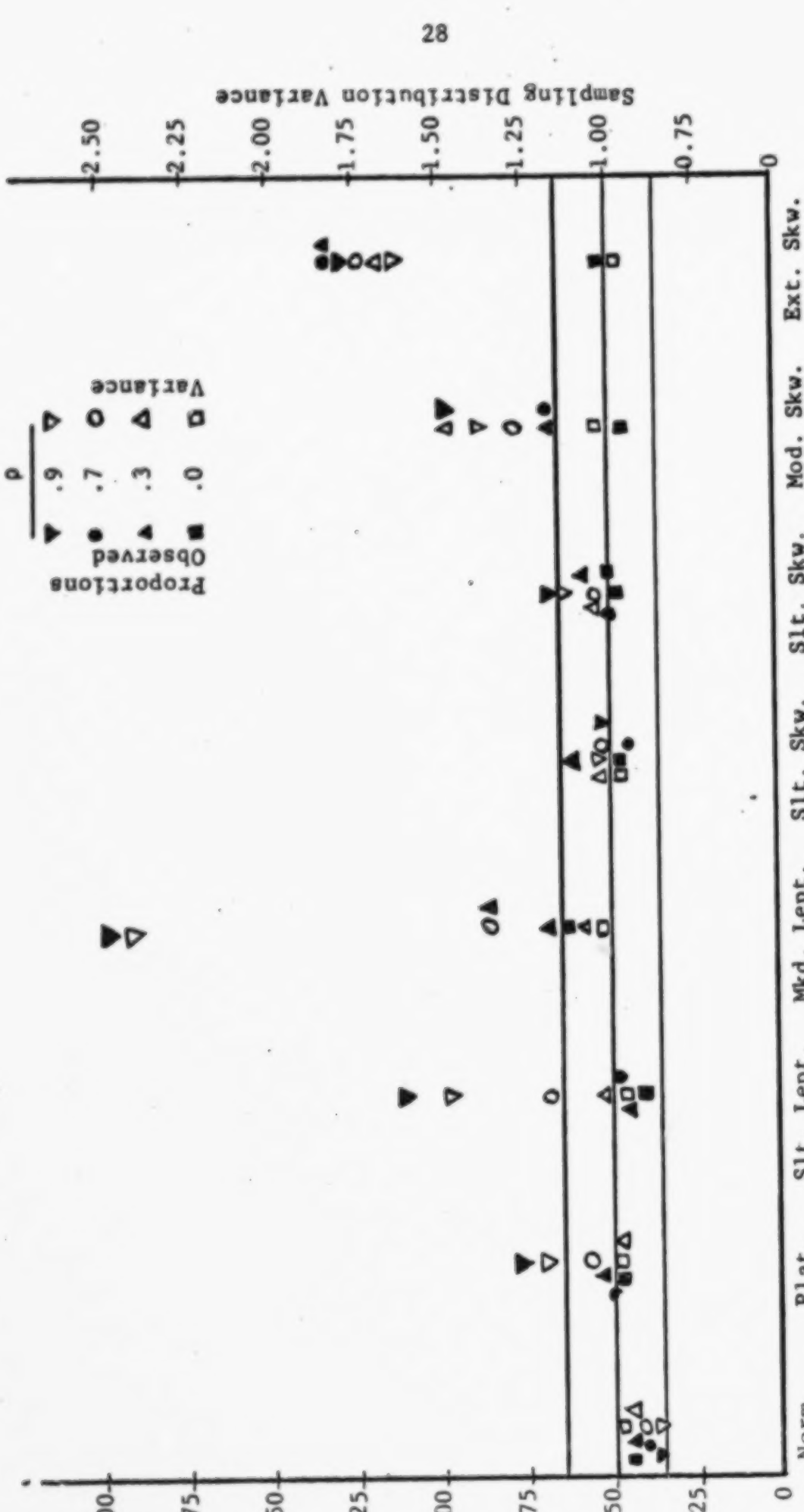

Figure 12. Proportions Observed in the Tails of the Sampling Distributions of $z_1 - z_2$ at the .05 Level of Significance, and Variances of the Sampling Distributions of $z_1 - z_2$, for All Population Types, Across All Levels of ρ, for Sample Size 100.

TABLE 6

OBSERVED VARIANCES OF THE SAMPLING DISTRIBUTIONS OF $z_1 - z_2$ AND $z^*_1 - z^*_2$ FOR ALL POPULATION TYPES, ACROSS ALL LEVELS OF ρ, FOR SAMPLES OF SIZE 100.

Treatment	ρ	$\sigma^2_{(z_1-z_2)}$	$\sigma^2_{(z^{**}_1-z^{**}_2)}$
Normal	.93	0.889	0.893
	.69	0.913	0.915
	.31	0.944	0.944
	.02	0.945	0.945
Platykurtic	.92	1.197	1.203
	.68	1.061	1.063
	.31	0.988	0.989
	.04	0.972	0.972
Slight Leptokurtic	.91	1.478	1.484
	.67	1.140	1.143
	.34	1.009	1.010
	.00	0.950	0.951
Marked Leptokurtic	.91	2.426	2.437
	.67	1.140	1.143
	.34	1.009	1.010
	.00	0.950	0.951
Slight Skew (Platy.)	.93	1.042	1.047
	.69	1.020	1.023
	.31	1.020	1.021
	.03	0.979	0.979
Slight Skew	.93	1.148	1.153
	.70	1.041	1.044
	.33	1.037	1.038
	.06	0.998	0.998
Moderate Skew	.92	1.383	1.390
	.68	1.286	1.289
	.31	1.146	1.147
	.06	1.023	1.024
Extreme Skew	.86	1.631	1.637
	.69	1.708	1.713
	.31	1.671	1.672
	.05	0.957	0.957

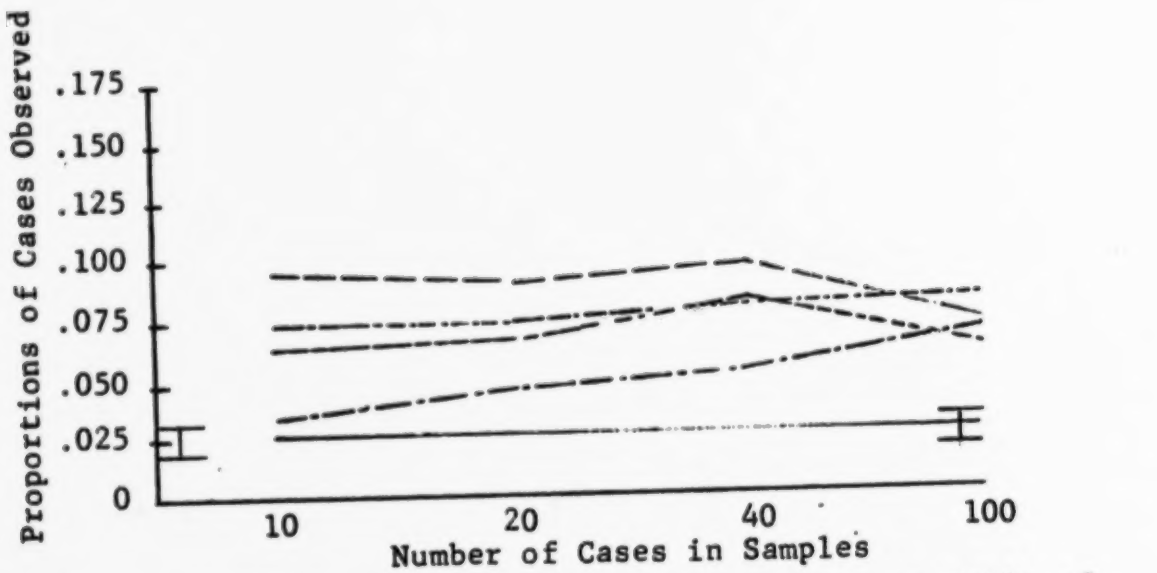

Figure 13. Proportions Observed in the Upper and Lower Tails of the Sampling Distributions of $\underline{z} - \zeta$ and $\underline{z}^* - \zeta^*$ at the .025 Level of Significance From an Extreme Skew Population Across All Sample Sizes, for ρ Approximately .90

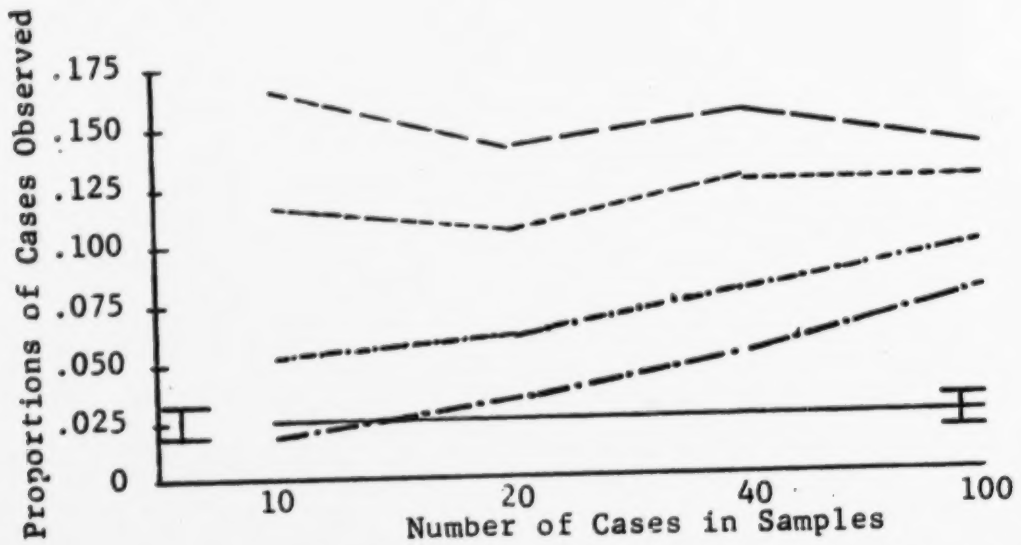

Figure 14. Proportions Observed in the Upper and Lower Tails of the Sampling Distributions of $\underline{z} - \zeta$ and $\underline{z}^* - \zeta^*$ at the .025 Level of Significance From a Marked Leptokurtic Population Across All Sample Sizes, for ρ Approximately .90.

The "mean r" and "mean r^*" of the sampling distributions of z and z^* are presented in columns two and three of Table 7. The values given for "mean r" and "mean r^*" are the derived values of r equivalent to the mean z and mean z^* of each sampling distribution. In the first column of Table 7 the values of ρ are given to four places. Comparison of the "mean r" and mean r^*" values with the actual population values reveals a most interesting finding. The "mean r^*" is closer than "mean r" is closer to the value of ρ, and for 6 populations there is no difference.

In all sampling distributions the upper tail-lower tail agreement is closer for the test ($z - \zeta$ or $z^* - \zeta^*$) having the closer agreement between "mean r" and ρ. Also the relative magnitude of proportions observed in the upper and lower tails is found to depend on the sign of (ρ - "mean r"). When "mean r" is less than ρ, greater proportions are observed in the lower tail, and when "mean r" is greater than ρ, greater proportions are observed in the upper tail. These relationships between "mean r," "mean r^*," and upper-lower tail proportions can be seen in one compares the values in Table 7 with the proportions observed plotted in Figures 13 and 14. No exceptions to the relationships described above are found in this study.

The observed measures of skewness, β_1, and kurtosis, β_2, are also given in Table 7 for the sampling distributions of z and z^* for samples of 100 cases. Comparison of β_1 and β_2 measures across z and z^* indicate that both distributions are very nearly normal across populations for all values of ρ. Based on the probability limits of β_1 and β_2 given by K. Pearson (1931) only four populations give rise to sampling distributions having less than a .05 probability of chance occurrence, and only two with probability of chance occurrence less than .01. All these departures are

TABLE 7.

OBSERVED VALUES OF "MEAN $\underline{r}$," "MEAN $\underline{r}^*$," β_1 AND β_2, FOR THE SAMPLING DISTRIBUTIONS OF $\underline{z}$ AND $\underline{z}^{**}$ FOR ALL POPULATION TYPES, ACROSS ALL LEVELS OF ρ, FOR SAMPLES OF SIZE 100.

Population Types	ρ	"Mean $\underline{r}$"	"Mean $\underline{r}^*$"	z		z*	
				β_1	β_2	β_1	β_2
Normal	.9312	.932	.930	0.00	2.93	0.00	2.93
	.6931	.969	.962	0.00	2.98	0.00	2.98
	.3123	.312	.309	0.00	2.90	0.00	2.90
	.0207	.022	.022	0.00	3.14	0.00	3.14
Platykurtic	.9260	.938	.926	0.01	3.10	0.01	3.10
	.6835	.686	.682	0.00	3.09	0.00	3.09
	.3127	.318	.314	0.00	2.91	0.00	2.91
	.0390	.036	.036	0.00	3.11	0.00	3.11
Slight Leptokurtic	.9052	.907	.905	0.03	3.06	0.03	3.06
	.6744	.676	.671	0.00	2.91	0.00	2.91
	.3386	.395	.341	0.01	3.15	0.01	3.15
	.0028	.001	.000	0.01	3.16	0.01	3.17
Marked Leptokurtic	.9117	.916	.913	0.06	3.12	0.06	3.12
	.6935	.697	.693	0.06	3.19*	0.06	3.19*
	.3200	.327	.323	0.08	3.14	0.08	3.14
	.0671	.070	.070	0.01	3.06	0.01	3.06

32

TABLE 7. Continued.

Population Types	ρ	"Mean r"	"Mean r*"	z		z*	
				β_1	β_2	β_1	β_2
Slight Skew (Platy.)	.9298	.931	.929	0.01	3.04	0.01	3.04
	.6905	.692	.688	0.00	2.90	0.00	2.90
	.3135	.319	.315	0.00	2.89	0.00	2.89
	.0313	.030	.030	0.00	3.24*	0.00	3.24
Slight Skew	.9305	.932	.930	0.01	3.08	0.01	3.08
	.6973	.699	.695	0.00	2.89	0.00	2.89
	.3318	.337	.334	0.00	2.92	0.00	2.92
	.0572	.057	.056	0.01	3.29**	0.01	3.29**
Moderate Skew	.9247	.926	.924	0.02	3.01	0.02	3.01
	.6794	.682	.678	0.00	2.88	0.00	2.88
	.3141	.319	.316	0.01	2.98	0.01	2.98
	.0563	.056	.056	0.03	3.33**	0.03	3.33**
Extreme Skew	.8607	.862	.859	0.00	3.05	0.00	3.05
	.6923	.694	.690	0.00	2.91	0.00	2.91
	.3103	.305	.302	0.03	2.92	0.03	2.92
	.0505	.054	.054	0.07	3.18	0.07	3.18

* Outside the .95 confidence interval.
** Outside the .99 confidence interval.

33

observed for β_2 and all indicate leptokurtic sampling distributions. Three of these β_2 values, and both the more disparate values, occur for ρ approximately .00. It should be noted that none of the β_2 values which fall outside the confidence intervals occur for populations having large excesses of proportions observed. This is in direct contrast to the variances of the sampling distributions which were found to be most disparate for those populations having large excesses of proportions observed. The values of β_1 and β_2 are identical for z and z^* (and for z^{**} as well, not shown in Table 7).

Summary of Results

For two-tailed tests of the form $z - \zeta$ the summary statements for tests of the form $z_1 - z_2$, the following summary statements seem to be warrented.

1. Within the range investigated (10 to 100), sample size does not appear to have any noticeable systematic effect on the sampling distribution of two-tailed tests of the form $z - \zeta$ and $z_1 - z_2$. (See Figure 9.)

2. For two-tailed tests of the form $z - \zeta$and $z_1 - z_2$ the results based on z and z^* are nearly identical. z^* and z^{**} differ only by the third term in the expression for z^{**}. Apparently the effect of this third term is negligible. (See Figure 10.)

3. For two-tailed tests of the form $z - \zeta$ and $z_1 - z_2$ it was observed that as ρ increases, the effects of marginal nonnormality of the bivariate distribution become more pronounced. This increased sensitivity to nonnormality is seen for skewed distributions even at ρ approximately equal to .30. ρ about .70 seems to mark the beginning of a rapid increase in sensitivity to leptokurtic distribution. (See Figure 11.)

4. For two-tailed tests of the form $\underline{z} - \zeta$ and $\underline{z}_1 - z_2$ the observed proportion excess introduced by marginal nonnormality seems to result from the effects of nonnormality upon the variance of the sampling distribution. Not only is a striking similarity seen between plots of variance observed and proportions obtained in the critical regions, but there seems to be little if any relationship between nonnormality of sampling distributions, as revealed by β_1 and β_2, and proportions obtained in the critical regions. Marginal nonnormality in the population seems to affect the variance rather than the normality of these sampling distributions. (See Figure 12.)

The following additional summary points are based on results observed for tests of the form $\underline{z} - \zeta$ when the tails of the sampling distributions were considered separately, i.e., for one-tailed tests. The results observed from investigation of the distribution characteristics of the sampling distributions of $\underline{z} - \zeta$, $\underline{z}$, $\underline{z}^*$, and $\underline{z}^{**}$ are also summarized below.

5. For tests of $\underline{z} - \zeta$ the proportions observed in the upper and lower tails of the sampling distributions are not equally distributed in the two tails. This inequality between the upper and lower tails is in all cases due to displacement of the observed mean of the sampling distribution relative to the expected mean. The upper tail ($\underline{z}$ greater than ζ) consistently contains a greater proportion of the observed cases than does the lower tail ($\underline{z}$ less than ζ). As the sample size is increased, the difference between the proportions observed in the two tails becomes smaller. (See Figures 13 and 14.)

6. The proportions observed in the upper and lower tails are more nearly equal for tests of $\underline{z}^* - \zeta^*$ than for $\underline{z} - \zeta$. For $\underline{z}^* - \zeta^*$, the upper tail-lower tail differences occur in both directions rather than appearing always in one direction, as is the case for $\underline{z} - \zeta$. These differences are not seen for the

total proportions observed in both tails. Upper tail-lower tail agreement is influenced by sample size, ρ, and population type.

7. The observed variance of the sampling distributions of $\underline{z}^* - \zeta^*$ and $\underline{z}^{**} - \zeta^{**}$ are consistently slightly greater than those for $\underline{z} - \zeta$. For ρ small in slightly nonnormal populations and in all normal populations $\underline{z}^* - \zeta^*$ and $\underline{z}^{**} - \zeta^{**}$ have sampling distribution variances slightly nearer the expected value of 1.0 than those for $\underline{z} - \zeta$.

8. The observed values of β_1 and β_2 for the sampling distributions of $\underline{z}$, $\underline{z}^*$, and $\underline{z}^{**}$ indicate that the normalizing effects of these transformations are identical. Most values of β_1 and β_2 are very close to the normal theory expected values. All distributions were symmetrical. There appears to be a tendency toward leptokurtic sampling distributions for ρ approximately equal to .00 in nonnormal populations. (See Table 7.)

9. The means of the sampling distributions of $\underline{z}^*$ tend to be nearer the expected means of the distributions than are the means of $\underline{z}$. As the differences between expected and observed sampling distribution means decrease, the differences between the proportions observed in the two tails of the distributions decrease. (See Table 7)

Implications for Practical Applications of z-Based Tests of r.

The above results indicate disturbances in the sampling distributions of sufficient magnitude to be of concern to most researchers. Yet the findings of this study offer no help to the research who is, and chooses to remain, ignorant of the general nature of the population he has sampled. For such a researcher, should one exist, there is in this study no basis for consoling statements concerning the robustness of z-based tests of r.

For the researcher who, by virtue of his experience within an area, has some estimate of the nature of the population he has sampled, two suggestions can be offered. First, if it is known that the population is nonnormal, one could consider the findings in this study for the population type most like the population in question. If β_1 and β_2 are unknown, comparisons with the marginal distributions shown in Figures 1 through 8 can be used to determine the most appropriate population. Based on the proportions observed for the most appropriate population type one could either adjust the significance level of tests or have an indication of the effective significance level in such a case. This is admittedly a crude technique requiring estimates which may be rather imprecise. Barring this approach, the results found suggest caution coupled with an awareness of the magnitude of error possible when dealing with nonnormal populations. Within the range and extent of marginal nonnormality investigated in the present study, the effective significance level for a chosen level of .05 could be as large as .25. For large ρ and marginal nonnormality an effective level of .15 to .20 for a chosen level of .05 cannot be considered uncommon.

The second suggestion is more easily implemented. If there is doubt as to the normality of the marginal distributions, especially if ρ is not small, it is suggested that z^* and z^{**} be employed rather than z. While

the use of z^* and z^{**} will not rule out the possibility of having an effective significance level considerable disparate from the chosen level, it will provide a much better balance of effective significance levels for the two tails of the test.

An additional note of caution concerning sample sizes seems justified. Except for the bias in the mean of z which can be largely avoided by use of z^* and z^{**}, the effects of nonnormality do not decrease very rapidly with increase in sample size. With the range of sample size studied (10 to 100), some effects were more pronounced for larger samples.

REFERENCES

Baker, George A. (1930), "The Significance of the Product-Moment Coefficient of Correlation with Special Reference to the Character of the Marginal Distributions," Journal of the American Statistical Association, 25, 287-96.

Chesire, Leone; Oldis, Elena; and Pearson, Egon (1932), "Further Experiments on the Sampling Distribution of the Correlation Coefficient," Journal of American Statistical Association, 37, 121-28.

David, F.N. (1938), Tables of the Distribution of the Correlation Coefficient. London: Biometrika Office.

Fisher, R. A. (1915), "Frequency Distribution of the Values of the Correlation Coefficient in Samples from an Indefinitely Large Population," Biometrika, 10, 507-21.

Fisher, R. A. (1921), "On the 'Probable Error' of a Coefficient of Correlation Deduced from a Small Sample," Metron, 1, 1-32.

Fisher, R. A. (1950), Statistical Methods for Research Workers, Eleventh Edition. London: Oliver and Boyd.

Games, P. A. and Lucas, P. A. (1966), "Power of the Analysis of Variance of Independent Groups on Non-Normal and Normally Transformed Data," Educational and Psychological Measurement, 26, (2), 311-27.

Gayen, A. K. (1951), "The Frequency Distribution of the Product-Moment Correlation Coefficient in Random Samples of Any Size Drawn from Non-Normal Universes," Biometrika, 38, 219-47.

Heath, H. A. (1961), "An Empirical Study of Correlation Involving a Half-Normal Distribution," Psychological Reports, 9, 95-86.

Hjelm, H. F. and Norris, R. C. (1962), "Empirical Study of the Efficacy of Fisher's z-Transformation," Journal of Experimental Education, 30, (3), 269-77.

Hotelling, H. (1953), "New Light on the Correlation Coefficient and Its Transforms," Journal of the Royal Statistical Society, Section B, 15, 193-225.

Knoble, H. D. (1968), "PRBA--Normal Probability Density Function (I)," Computer Program,Fortran IV, The Pennsylvania State University, Computation Center.

Norris, R. C. and Hjelm, H. F. (1961), "Nonnormality and Product Moment Correlation," Journal Experimental Education, 29, (3), 261-70.

Norton, D. W. (1952), "An Empirical Investigation of Some Effects of Non-normality and Heterogeneity on the F-Distribution," Ph.D. Dissertation in Education, State University of Iowa. Cited in Lindquist, E. F. Design and Analysis of Experiments in Psychology and Education. Boston: Houghton Mifflin Company, 1956. 78-90.

Pearson, Egon S. (1931), "The Test of Significance for the Correlation Coefficient," Journal of the American Statistical Association, 26, 123-24.

Pearson, Egon S. (1932), "The Test of Significance for the Correlation Coefficient: Some Further Results," Journal of the American Statistical Association, 27, 424-26.

Rider, P. R. (1932), "On the Distribution of the Correlation Coefficient in Small Samples," Biometrika, 24, 382-403.

Soper, H. E.; Young, A. W.; Cave, B. M.; Lee, A.; and Pearson, K. (1916), "On the Distribution of the Correlation Coefficient in Small Samples. Appendix 11 to the Papers of 'Student' and R. A. Fisher. A Cooperative Study," Biometrika, 11, 328-413. (Cited as the Cooperative Study.)

'Student,' (1908), "Probable Error of a Correlation Coefficient," Biometrika, 6, 302-10.